Ludwig-Maximilians-Universität München
Institut für Informatik
PD Dr. Matthias Schubert
Markus Mauder

Knowledge Discovery in Databases II
WS 2014/2015

Übungsblatt 11: Graph and Link Mining

Aufgabe 11-1   Graphlet Kernels

Graphlets are a means to transform the topology of a graph into a vector space. In this context, a graphlet is a sub-graph with exactly k vertices. k-graphlets differ from each other in topology, i.e. the configuration of vertices. The graph $G(V, E)$ is expressed as a vector through the number of individual k graphlets in $G(V, E)$.

(a) How many graphlets of size 4 does a graph with n vertices contain? What is the time complexity of extracting them?

(b) How expensive is comparing two graphs using graphlets of size 4?

(c) How can graphlet vector comparison be sped up/improved?

(d) What problems emerge with larger values of k? What are the consequences of using an vertex-labeled graph $G(V, E, L)$?

Aufgabe 11-2   Shortest Path Kernel and Wiener Index

Compare two graphs using Shortest Path Kernel.

Additionally, let the Wiener Index $W(G)$ on a graph G be defined as:

$$W(G) = \sum_{v_i \in G} \sum_{v_j \in G} d(v_i, v_j),$$

with $d(v_i, v_j)$ the shortest path length between vertices v_i and v_j from G.

What is the connection between the Wiener Index and the Shortest Path Kernel?